ICHA 2018

HARMFUL ALGAE 2018 – FROM ECOSYSTEMS TO SOCIO-ECOSYSTEMS

PROCEEDINGS OF THE 18TH INTERNATIONAL CONFERENCE ON HARMFUL ALGAE

21-26 October 2018, Nantes, France

International Society for the Study of Harmful Algae
ISSHA
Ifremer
UNESCO
United Nations Educational, Scientific and Cultural Organization
Intergovernmental Oceanographic Commission

ISBN 978-87-990827-7-3

HARMFUL ALGAE 2018 – FROM ECOSYSTEMS TO SOCIO-ECOSYSTEMS

PROCEEDINGS OF THE 18TH INTERNATIONAL CONFERENCE ON HARMFUL ALGAE

21-26 October 2018, Nantes, France

Editor: Philipp Hess (Ifremer)

Graphic design : MCI France
Lay-out : Hélène Parfait (Ifremer)

Published by the International Society for the Study of Harmful Algae (ISSHA) and
the Institut Francais de Recherche pour l'Exploitation de la Mer (Ifremer),
in cooperation with the Intergovernmental Oceanographic Commission of the United Nations Educational, Scientific and Cultural Organization (IOC/UNESCO).

For bibliographic purposes, this document should be cited as follows:

Ph. Hess [Ed]. 2020. *Harmful Algae 2018 – from ecosystems to socioecosystems. Proceedings of the 18th Intl. Conf. on Harmful Algae*. Nantes. International Society for the Study of Harmful Algae. 214 pages. ISBN: 978-87-990827-7-3.

DISCLAIMER

ICHA 2018

21 – 26 OCTOBER

NANTES, FRANCE

THE 18TH INTERNATIONAL CONFERENCE

ON HARMFUL ALGAE

Hosted by

Institut Français de Recherche pour l'Exploitation de la Mer (Ifremer)

Convener

Philipp Hess, Ifremer Centre Atlantique, Nantes, France

Under the auspices of

The International Society for the Study of Harmful Algae (ISSHA)

COMMITTEES

Local organizing committee

Anthony MASSE
Université de Nantes, France

Gaël BOUGARAN,
Ifremer, France

Hélène PARFAIT
Ifremer, France

Maud LEMOINE
Ifremer, France

Patricia THIBAULT-PRALIN
Ifremer, France

Philipp HESS
Ifremer, France

Sophie PILVEN
Ifremer, France

Véronique SECHET
Ifremer, France

Zouher AMZIL
Ifremer, France

Scientific Committee (reviewers underlined)

Allan CEMBELLA
Alfred-Wegener-Institut, Germany

Amel LATIFI
Aix Marseille Université, France

Anne THEBAULT
Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail, France

Annick WILMOTTE
Université de Liège, Belgium

Anthony MASSE
Université de Nantes, France

Benjamin MARIE
Muséum National d'Histoire Naturelle, France

Catherine BELIN
Ifremer, France

Catherine DREANNO
Ifremer, France

Catherine QUIBLIER
Museum National d'Histoire Naturelle, France

Cécile BERNARD
Museum National d'Histoire Naturelle, France

Claudia WIEGAND
Université de Rennes 1, France

Clémence GATTI
Institut Louis Malardé, France

Damien TRAN
Université de Bordeaux, France

Delphine LATOUR
Université Blaise-Pascal de Clermont-Ferrand, France

Felipe ARTIGAS
Université Littoral Côte d'Opale, France

Gaël BOUGARAN
Ifremer, France

Hélène HÉGARET
Université de Brest, France

Jean-François HUMBERT
Université Pierre et Marie Curie, France

Jean-Luc ROLLAND
Ifremer, France

Jordi MOLGO
Institut des Neurosciences Paris-Saclay, France

Juliette FAUCHOT
Université de Caen, France

Laure GUILLOU
Station biologique de Roscoff, France

Marc SOURISSEAU
Ifremer, France

Marie BONNIN
Institut de Recherche pour le Développement, France

Marie-Yasmine DECHRAOUI-BOTTEIN
International Atomic Energy Agency, Monaco

Mathilde PALUSSIERE
DGAl/Bureau des Produits de la Mer et de l'Eau Douce, France

Maud LEMOINE
Ifremer, France

Mélanie ROUE
Institut de Recherche pour le Développement, France

Mireille CHINAIN
Institut Louis Malardé, France

Mohamed LAABIR
Université de Montpellier, France

Muriel GUGGER
Institut Pasteur, France

Myriam BORMANS
Université de Rennes 1, France

Nathalie ARNICH
Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail, France

Nicolas CHOMERAT
Ifremer, France

Philipp HESS
Ifremer, France

Philippe SOUDANT
Université de Brest, France

Raffaele SIANO
Ifremer, France

Rodolphe LEMÉE
Observatoire océanologique de Villefranche sur Mer, France

Romulo ARAOZ
Commissariat à l'Energie Atomique, France

Ronel BIRÉ
Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail, France

Sophie PARDO
Université de Nantes, France

Valérie FESSARD
Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail, France

Véronique MARTIN-JEZEQUEL
Université de Nantes, France

Véronique SECHET
Ifremer, France

Zouher AMZIL
Ifremer, France

International Advisory Board (reviewers underlined)

Aifeng LI
China

Allan CEMBELLA
Germany

Ana GAGO
Spain

Asma HAMZA
Tunisia

Bastiaan Willem IBELINGS
Switzerland, Netherlands

Beatriz REGUERA*
Spain

Charles LUGOMELA
Tanzania

Christine BAND-SCHMITT
Mexico

Don ANDERSON
USA

Emanuela TESTAI
Italy

Esther GARCÈS*
Spain

Grant PITCHER
South Africa

Gustaaf HALLEGRAEF*
Australia

Henrik ENEVOLDSEN*
Denmark

Ian JENKINSON*
China, France

Lincoln MACKENZIE*
New Zealand

Lirong SONG
China

Luis BOTANA
Spain

Luis PROENCA
Brazil

Marie-Yasmine DECHRAOUI-BOTTEIN
Monaco

Mitsunori IWATAKI
Japan

Patricia TESTER
USA

Philipp HESS*
France

Rainer KURMAYER
Austria

Rencheng YU
China

Robin RAINE*
Ireland

Sandra SHUMWAY
USA

Songhui LU
China

Sun Ju KIM
Republic of Korea

Takeshi YASUMOTO
Japan

Triantafyllos KALOUDIS
Greece

Vera TRAINER*
USA

**members of the ISSHA Council Conference Organizing Sub-committee*

ADDITIONAL PROCEEDINGS REVIEWERS

Adriana ZINGONE
Stazione Zoologica Anton Dohrn Napoli, Italy

Antonella PENNA
Università degli Studi di Urbino Carlo Bo, Italy

Bengt KARLSON
Swedish Meteorological and Hydrological Institute, Sweden

Elisa BERDALET
Institut de Ciències del Mar, Spain

Fran VAN DOLAH
National Oceanic and Atmospheric Administration, USA

Keith DAVIDSON
Scottish Association for Marine Science, Scotland

Leila BASTI
Tokyo University of Marine Science and Technology, Japan

Luiz MAFRA
Universidade Federal do Paraná, Brazil

ØJvind MOESTRUP
University of Copenhagen, Denmark

Patrice GUILLOTREAU
Université de Nantes, France

Po Teen LIM
University of Malaya, Malaysia

Raphaele LE GARREC
Université de Bretagne Occidentale, France

Raphe KUDELA
University of California, Santa Cruz, USA

Robert HATFIELD
Centre for Environment, Fisheries and Aquaculture Science, UK

Urban TILLMANN
Alfred Wegener Institute, Germany

Uwe JOHN
Alfred Wegener Institute, Germany

Wayne LITAKER
National Oceanic & Atmospheric Administration, USA

PARTNERSHIPS

Bronze sponsors

French Research Project ANR OSS-Cyano
Marine Drugs
McLane
Phytoxigene
Synoxis Algae

Other sponsors

Cambridge Isotope Laboratories
Cifga
Cytobuoy B.V.
Fluid Imaging Technologies
IKA Werke GmbH & Co. KG
Jerico-Next
Marine Institute
Merck KGaA
Novakits Abraxis
Photon Systems Instruments
Statera Environmental
Station Ifremer Concarneau

French Institutional Support

Institut Français de Recherche pour l'Exploitation de la Mer (Ifremer)
GdR Phycotox (Ifremer – CNRS 3569)
Agence Française pour la Biodiversité (AFB)
Commissariat à l'énergie atomique et aux énergies alternatives (CEA)
Museum National d'Histoires Naturelles
Nantes Métropole
Université Bretagne Loire
Institut Universitaire Mer et Littoral (IUML)

International Institutional Cooperation

International Atomic Energy Agency (IAEA)
U.S. National Institute of Environmental Health Sciences (NIEHS)
Scientific Committee on Oceanic Research (SCOR)
U.S. National Oceanic and Atmospheric Administration (NOAA)

EDITOR'S PREFACE

The 18th International Conference on Harmful Algae (ICHA) was held at La Cité des Congrès, Nantes, France, 21-26 October 2018. After the last conference in the same spot 25 years earlier, this was the second time the conference was convened in France. The conference was hosted by Ifremer and was born from an initiative of the French marine HAB research community (GdR PHYCOTOX), with strong support by its sister organization for freshwater cyanobacteria (GIS Cyano). The theme of the conference "From Ecosystems to Socioecosystems" was chosen to stress the need to integrate natural sciences with social sciences to increase the impact of our science on society. A short summary of the event and a scientific summary have already been published in *Harmful Algae News* (pages 1-8, issue 62; June, 2019).

Compared to previous conferences, this edition has enjoyed record attendance (709 participants from 64 countries) and scientific contributions (613 abstracts, with 255 oral presentations and 358 posters, 45 of which were also presented as ignite talks). Out of nine plenary lectures, five were given by scientists not usually attending this conference. This is a testimony to our efforts in soliciting scientists from adjacent scientific fields. They gave us insights into recent developments regarding linkages between climate change and eutrophication (Anna Michalak), in human toxicology (Thomas Hartung), trait-based ecology (Elena Litchman), natural products chemistry (William Gerwick) and remediation (Eric Jeppesen). The four plenaries directly related to HAB science focused on HAB parasites (Laure Guillou), Ciguatera (Mireille Chinain), chemical ecology (Allan Cembella, Yasumoto award lecture 1) and analytical measurements (Michael Quilliam, Yasumoto award lecture 2). The other lectures, like the posters, were subdivided into 21 themes, 45 oral sessions and three lunch time seminars. Credit for this scientifically diverse and inspiring program largely goes to our national scientific committee built on GdR Phycotox and GIS Cyano (43 people). Overall feedback on this ambitious program was very positive (87 participants (12%) replied to our feedback questionnaire: 80% of the respondents found the scientific program good or excellent). We heavily drew on the ISSHA council, our national scientific committee and other key scientific members of our community for our 53 session chairs – again, thank you all for helping us in this event!

I also acknowledge the help of our reviewers of the 47 submissions to the proceedings (> 90% reviewed by 2 referees). A frustrating aspect of conference proceedings is the change in publication attitude of conference attendees. While most participants contributed to the conference proceedings in 1989 (100 contributions), less than 8% of contributions to the conference in 2018 resulted in 47 submissions to the proceedings. This lack in participation to the proceedings obviously results from the need for "measurable outputs" in form of peer-reviewed papers in journals indexed in the "Web of Science". Still, it also means that conference proceedings no longer actually reflect the science presented at the conference. For a comprehensive overview of the conference, please download the final programme or the abstract book.

Participation and number of contributions has approximately doubled since the last conference in Nantes in 1993. The fourth *International Conference on Toxic Marine Phytoplankton* (26-30 June 1989, Lund, Sweden) was the first one to extend coverage beyond dinoflagellates to all kinds of marine microalgae, including benthic or epiphytic species. Since that time, the breadth of the conference has gradually increased which partly explains the increased participation. For instance, the 2018 edition also had a very strong component on cyanobacterial ecology, toxins and remediation. Another interesting example is the report of the structure of Ciguatoxin (= CTX1B = P-CTX1) by

Takeshi Yasumoto, which was one of the five contributions on *Gambierdiscus* and its toxins in 1989; comparatively, there were 49 contributions to this theme at the conference 30 years later, which is in line with recent IOC priorization of this problem. Similarly, 7 contributions dealt with effects of HABs on aquatic organisms in 1989, shortly after a major bloom of *Chrysochromulina polylepsis* in the Skagerrat, while 23 contributions dealt with this theme in 2018, indicating an increased interest in this topic over time. The most striking paradox comparing these two conferences is that the 1989 event finishes its overview (Max Taylor, page 532 of the proceedings) with a major question dealt with during the conference: Are HABs increasing and are they spreading? As outlined by Gustaaf Hallegraeff in its HAN article (HAN 62, pages 9-11), 139 abstracts of the 2018 Nantes meeting dealt with the same question, formulated in the context of climate change. Thus, three major topics in the 2020s (ciguatera, fish kills and evolution of HABs in the context of climate change) have been subject of discussion by our community for several decades. However, there were also many more recent developments discussed at the conference, drawing on modern techniques. The three lunch time seminars were dedicated to (i) sensors and drones for surveillance of freshwater HABs, (ii) *in-situ* monitoring tools for marine HAB species, such as the *Imaging Flow CytoBot*, and (iii) molecular biology techniques such as qPCR for monitoring of species difficult to identify with light microscopy. These seminars were a clear testimony to our increased capacity to detect and monitor an ever wider range of species more efficiently. Recent advances in the Omics field (43 contributions) also led to significant progress reported at this conference on toxin discovery (> 100 prymnesins & cyanotoxins), our understanding in toxin biosynthesis (domoic acid & polyketides) and interactions of HAB species with aquatic animals (copepodamides).

Overall, the conference was also financially viable since income provided by participants (72 %) and sponsors (20 %) covered 92 % of total costs incurred (455k€), while Ifremer recovered all direct expenses, excluding staff time. Thanks also go to Hélène Parfait, the secretary of the Phycotoxins Laboratory (Ifremer, Nantes), and Cécile Salaun (administrative project officer, Ifremer, Brest) who both accompanied me in managing the professional conference organizer and budgets on a daily basis. There was a reasonable equilibrium between public (14) and private (12) sponsors, who I thank for balancing our budget. Many, many thanks also to our Phycotoxins team in Nantes who lovingly helped during countless hours with logistics during the preparation and the event itself. Finally, I'd like to express my gratitude to our directorate at Ifremer for taking on the financial risk of this undertaking as early as 2014, at a time when nothing was known of Zika-virus, French protests against carbon taxes (yellow vests) or Corona-virus, to name but some of the risks conference organizers have to deal with. We wish our Mexican and Central American colleagues all the best of courage, patience and luck for the organization of the 19th ICHA in La Paz, which will be held in 2021 due to latest developments.

TABLE OF CONTENTS

Combination of machine learning methodologies and imaging-in-flow systems for the automated detection of Harmful Algae

Guillaume Wacquet[1,2,3,*], Alain Lefebvre[2], Camille Blondel[2], Arnaud Louchart[1], Philippe Grosjean[3], Nadine Neaud-Masson[4], Catherine Belin[4], Luis Felipe Artigas[1,*]

[1] Laboratory of Oceanology and Geosciences (UMR 8187 LOG), CNRS, ULCO, ULille, Wimereux, France
[2] Ifremer, Environment and Resources Laboratory (LER), Boulogne-sur-Mer, France
[3] Univ. Mons, Complexys Institute, Laboratory of Numerical Ecology of Aquatic Systems, Mons, Belgium
[4] Ifremer, Information System for Integrated Management and Monitoring (VIGIES), Nantes, France

[*] corresponding authors: guillaume.wacquet@univ-littoral.fr; felipe.artigas@univ-littoral.fr

Abstract

In recent years, improvements in data acquisition techniques have been carried out to sample, characterize and quantify phytoplankton communities at high temporal and geographical resolution, with a special focus on potential harmful algae, during oceanographic campaigns or in the frame of monitoring networks (to support knowledge but also for EU Directives and Regional Sea Convention needs). These acquisition and digitization techniques, including "imaging-in-flow" systems, allow to process a high number of samples and, consequently, generate an important quantity of data in which the presence of target events might not be detected. Indeed, as for traditional samples analysis with inverted microscope, a full manual quantification of the particles based on a simple visual inspection can be time-consuming, tedious and consequently lead to erroneous or wrong identifications.

For this purpose, the ZooImage R-package was and is still being developed to allow greater automation in data classification and analysis while also permitting some user-interaction during the process. The proposed methodology consists in combining few expert knowledge and machine learning algorithms at different levels: (i) to classify particles into different groups based on the definition and the adaptation of a specific training set through the use of "contextual data"; (ii) to detect and partially validate the "most suspect" predictions, based on a probability of misclassification; (iii) to estimate the number of cells for each colonial form thanks to the building of specific predictive models.

These different semi-automated tools were applied to the *in vivo* image dataset acquired with the FlowCam instrument during the September-October CAMANOC 2014 (Ifremer) cruise in the English Channel, in order to evaluate their operational ability to monitor the diversity of samples for the microphytoplankton, and especially to detect, track and count the most frequent potentially harmful algae found in this area at that period, like species belonging to *Pseudo-nitzschia*, *Dinophysis*, *Prorocentrum* and *Phaeocystis* genera. A distribution of these target groups was computed which highlights different sub-regions in the English Channel during the late summer-fall transition.

Keywords: Machine learning, User-interaction, Semi-automated classification, English Channel, HABs

Introduction

Since 2000, phytoplankton is defined as a biological parameter for marine quality assessment by the Water Framework Directive (WFD) and in 2008, the European Marine Strategy Framework Directive (MSFD) confirms this capacity. Today, about 7000 different species have been identified worldwide of which about 70 are potentially harmful (toxins or foam producers). Because of the consequences they have on the ecosystems goods and services, their economic and health impacts, the occurrence of these toxic events contributed to renewed interest towards studies on marine ecosystem, including ecological, climatic and economic domains.

The traditional technique to determine the phytoplankton composition of a sea water sample, is the identification and enumeration by inverted microscopy method after sedimentation (Utermöhl, 1958). However, this method is time consuming and requires experts specialized in phytoplankton

taxonomy. Moreover, it cannot detect spatial and temporal changes in the marine ecosystem combining high frequency and quality of information obtained without competent human participation to collect and classify these samples. Recent advances in image analysis and pattern recognition have made possible to estimate microphytoplankton concentration without hard numerical process, directly on living samples (Tang *et al.,* 1998). These approaches, using digital imaging of phytoplankton particles measured by image analysis and classified using machine learning algorithms, reduce the analysis time, improve the counting accuracy and can carry out a lot of measurement providing opportunities for microbiologists to get new insights into the functioning of communities in the nano-microplankton size range (Benfield *et al.,* 2007).

Over the past few decades, different specialized systems coupling digitization devices and data processing software have been developed for enumeration and measurements of particles in the nano-microplankton size range, like the Imaging FlowCytoBot instrument (IFCB) coupled to a specific data processing and classification software (Olson and Sosik, 2007), or the FlowCam® system (Fluid Imaging Technologies, Inc.[1]) associated to the VisualSpreadSheet® software (Sieracki *et al.*, 1998). Unfortunately, these software systems are designed for specific devices and are not always compatible with other imaging systems. Additionally, recent optimization of plankton imagers leads to faster digitization, multiplication of analyses and accumulation of a large quantity of data and images.

In this context, we propose to develop, to test and to prove the added-value of a semi-automated identification system for phytoplankton classification. This system could improve our knowledge of the ecology of phytoplankton by allowing semi-automated analyses to be applied to a higher temporal/spatial resolution than traditional techniques, due to time saving during data acquisition and processing steps (including validation of the results). These analyses are thus better adapted to study the frequencies of phenomena that control this biological compartment (Cloern, 1996). The proposed tool consists in coupling the FlowCam device for acquiring the digital images of phytoplankton particles in a sample, with the ZooImage R-package[2], which allows identifying and counting phytoplankton in a semi-automated way (Grosjean and Denis, 2013).

Materials and Methods

Sampling strategy

The CAMANOC cruise was a multidisciplinary survey within the framework of an ecosystem approach to fisheries. It was conducted from September 16 to October 12, 2014, in the Western and Central English Channel, on-board the R/V "Thalassa" (IFREMER). Water samples were collected at 89 sampling stations at subsurface (<0.5m), using a 5-L Niskin bottle. These samples were maintained cold (4°C) and in the dark until their analysis.

Imaging-in-flow system

Samples were digitized using an 8-bit grayscale benchtop FlowCam® VS with the pump speed set to 1.8 $ml.min^{-1}$. A 4X objective (40X overall magnification) coupled with a 300µm-depth flow-cell was used and samples were run in "AutoImage" operation mode. As described by Zarauz *et al.* (2007), all particles in the field of view of the camera (phytoplankton, zooplankton and inorganic particles) are imaged and captured at a regular user-defined interval, which allows an accurate estimation of imaged volume and consequently of particle concentration.

The specific software program, provided with the FlowCam device and named VisualSpreadSheet, is essential for all the major aspects of analysis: setup for data acquisition through the context settings for controlling the device, managing files and setting preferences, data acquisition and post-processing of collected data. For each particle, we obtain a set of 26 image parameters: 8 basic shape parameters (area, length, width, etc.), 13 advanced morphological parameters (circularity, convexity, roughness, etc.) and 5 grayscale measurements (intensity, transparency, etc.). The rest of the process, from image processing to statistical analysis, is done using the version 5.5.2 of the ZooImage R-package.

Training set and classifier

Building a training set is a crucial step in the automated recognition of plankton. Indeed, it represents the database used to generate a tool for automated or semi-automated recognition. Within this study, a training set representative of each plankton community met in the English Channel,

[1] https://www.fluidimaging.com/

[2] https://cran.r-project.org/web/packages/zooimage/

was built using samples taken throughout 2013 (in the frame of the Ifremer Regional Nutrients Monitoring network) and 2014 (in the frame of the CNRS LOG DYPHYRAD monitoring transect). A total of 3582 images were manually classified in 28 plankton groups. Moreover, instead of manually removing detrital particles and artefacts as is commonly done (Zarauz *et al.*, 2007), we added 12 groups for floating dark and light dead particles, bubbles, fibers, etc. to the 28 plankton groups in order to automatically identify and then eliminate them from the statistics. Finally, 5154 images were sorted into 40 groups. From this training set, a classifier is trained using the "Random Forest" algorithm (Breiman, 2001). Global error measured by using 10-folds cross-validation is equal to 25% for this training set.

Adaptive learning

Phytoplankton communities are sensitive to environmental and climate changes and modifications in their composition can occur at fast rates, which can lead to significant prediction errors when a static training set is used. Indeed, the representativeness and the variability of the particles into each group in the training set can have an important impact on the classification results. However, in order to build an ideal training set, a large number of labelled particles is needed, but also requires a lot of time, human resources and skills. To overcome this problem, adaptive learning improves a simple training set by adding labelled items taken from a pool of candidates.

This pool is composed of particles belonging to samples taken in the same experimental conditions than the studied one, and already validated by an expert. These "contextual" data allow to adapt the training set temporally and geographically to the phytoplankton communities generally encountered in the studied area and consequently to reduce the initial prediction error by 20-30%. In this study, one sample was selected every 2 days and all particles were manually labelled and validated. These particles were added to the initial training set in order to build an adapted recognition tool which is then used to classify the samples of the next 2 days.

Semi-automated validation

Performance scores obtained by completely automated classification are often considered too low to provide relevant abundances for all groups. Therefore, it is necessary to validate these predictions (Gorsky *et al.*, 2010). However, the validation of all predictions can be tedious and time-consuming if the sample size is important. In the ZooImage package, sophisticated techniques for semi-automated recognition are proposed. This allows a less biased estimation of abundance without having to manually validate all images.

In this context, the images which must be validated are detected by the automated classification system on the basis of a probability of a correct prediction for each particle. Once the images validated, the error can be modelled in order to perform statistical corrections. This approach offers an ideal trade-off between the full-automated method and the total manual validation while guaranteeing similar or better performances of recognition within an acceptable time. Moreover, the statistical correction of the remaining errors associated to the adaptive learning method, allows decreasing the global error more rapidly than with the initial training set.

Single cell counts

Usually, with automated acquisition systems like FlowCam, IFCB, flow cytometry, etc., colonies are considered and subsequently counted as one single particle, the same as single cells. Moreover, although colonies contribute largely to annual productivity, all biomass estimators are essentially calibrated on abundance in terms of single cells per volume unit. That is why the automated counting of the number of single cells in the various colonial forms represents a real challenge.

The proposed method consists in building specific predictive models for each studied group, based on the manual counts made on images contained in the training set. For this, a multivariate analysis was performed thanks to a simple linear modelling.

Results and Discussion

HABs distribution in the English Channel

An important spatial heterogeneity in microphytoplankton distribution was observed all along the sampling area. In order to highlight specific areas for the potential HABs and the associated spatial distribution, a clustering method was applied to their abundances. The Partitioning Around Medoids algorithm (PAM) showed 3 distinct areas (Figure 1):

- Western English Channel (WEC) where HABs were essentially characterized by *Pseudo-nitzschia* genus;

- Central English Channel (CEC) with high abundances of dinoflagellates (*Prorocentrum* genus) and haptophytes (*Phaeocystis* genus);
- Bay of Seine (BOS) where the highest abundances of *Pseudo-nitzschia* and *Prorocentrum* genera were observed.

Microscopy comparison and single cell counts

In order to validate these results, a comparison with the abundances obtained by microscopic counts was performed in some stations in the WEC, and especially for the most abundant genus which is *Pseudo-nitzschia*. As shown on Figure 2, the raw abundance (*i.e.* with colonial forms) estimates obtained with the FlowCam were always lower than counted with the inverted microscope, for which all cells were counted, whether they be in a colonial or a single cell form.

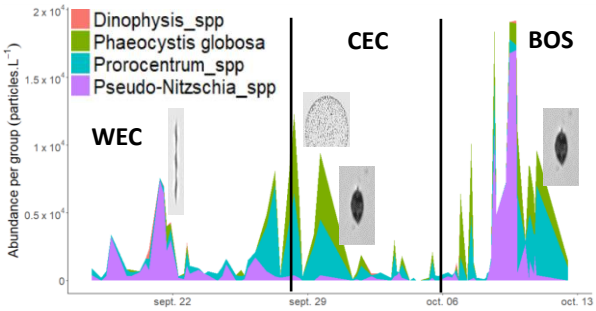

Fig. 1. Variability of the abundances (particles.L^{-1}) of each harmful genera during the CAMANOC cruise.

However, despite this under-estimation, a significant Spearman's rank correlation coefficient was found between *Pseudo-nitzschia* FlowCam and microscopic counts (Table 1).

After application of the specific predictive model for *Pseudo-nitzschia*, the overall dynamics of abundance obtained by microscopy and FlowCam were similar (Figure 2). In addition, an increase of the correlation coefficient was observed.

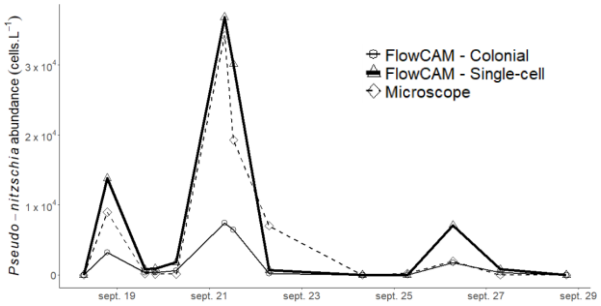

Fig. 2. Comparison between cell abundance obtained by microscopy (dashed line) and FlowCam with () and without single cell enumeration for *Pseudo-nitzschia* genus in some stations in WEC.

The results presented in this study highlight the need and the relevance of the combination of the automated process and the integration of some expert knowledge. However, in order to lead to more accurate discrimination, an optimization of training sets and algorithms, associated with the improvement of the resolution of the digital images, is still needed. Moreover, it is important to note that the ZooImage package, and consequently all the (semi-)automated tools presented in this study, can also be used on images acquired by various automated systems, like ZooScan, IFCB, etc.

Table 1. Spearman's rank correlation coefficient between *Pseudo-nitzschia* abundances obtained by microscopy and by FlowCam analysis, considering colonial forms or single cells counting.

Comparison	Spearman	p-value
Colonial form	0.74	< 0.05
Single cell	0.79	< 0.05

Acknowledgements

The authors thank the P.I. of the CAMANOC cruise (Morgane Travers-Trollet), as well as the captain and crew of R/V "Thalassa". Guillaume Wacquet was supported by post-doctoral grants funded by the Complexys Institute of the University of Mons and by the "Conseil Régional des Hauts-de-France" and is currently supported by the Centre National de la Recherche Scientifique (CNRS) in the frame of the MSFD convention (n°2101893310) with the French Ministry for an Ecological and Inclusive Transition (MTES), for the monitoring programme for pelagic habitats.

References

Benfield M. C., Grosjean Ph., Culverhouse P. F., Irigoien X., Sieracki M. E., Lopez-Urrutia A., Dam H. G., Hu Q., Davis C. S., Hansen A., Pilskaln C. H., Riseman E. M., Schultz H., Gorsky G. (2007). Oceanography 20, 172-187.

Breiman, L. B. (2001). Mach. Learn 45, 5-32.

Cloern, J.E. (1996). Geophysics 34(2), 127-168.

Gorsky G., Ohman M.D., Picheral M., Gasparini S., Stemmann L., Romagnan J.-B., Cawood A., Pesant S., García-Comas C., Prejger F. (2010). Plankton Research 32(3), 285–303.

Grosjean Ph., Denis K. (2013). Data Mining Applications with R, Elsevier, 978-0-12-411511-8, Chapter 12.

Olson R.J., Sosik H.M. (2007). Limnol. Oceanogr. 5, 195-203.

Sieracki C. K, Sieracki M. E., Yentsch C. S. (1998). Mar. Ecol. Prog. Ser. 168, 285-296.

Tang X., Stewart W. K., Vincent L. U. C., Huang H. E., Marra M., Gallager S. M., Davis C. S. (1998). Artificial Intelligence Review 12, 177-199.

Utermöhl H. (1958). Mitt. Int. Ver. Limnol. 9, 1-38.

Zarauz L., Irigoien X., Urtizberea A., Gonzalez M. (2007). Mar. Ecol. Prog. Ser. 345, 27–39.

Proceedings of the 18th International Conference on Harmful Algae

www.issha.org

ISBN 978-87-990827-7-3